ABSTRACT OF THE DISCLOSURE

In the similarity calculating portion, a similarity between frame feature vectors of the subject videos for retrieval, stored in the frame feature vector storing portion, and image feature vectors of the query images extracted by the image feature vector extracting portion is calculated, further in the frame feature vector selecting portion, the frame feature vectors of which the similarities are higher than a predetermined threshold value are selected, and furthermore, in the similar segment generating portion, the frame feature vectors that are continuous in time, out of those selected by the frame feature vector selecting portion, are integrated into one group. Thereafter, in the similar image selecting portion, at least one frame feature vector of the highest similarity inside the integrated group is selected so as to present an image having the selected frame feature vector as a result of retrieval.